

**Exercice 1*****Weary water*****8 points**

Vous venez d'arriver en stage, votre patron vous demande de l'aider à définir la stratégie de sa start up par rapport au problème de sélection de variables dans leur super module existant d'apprentissage statistique

1. Qu'est-ce que la sélection de variable et à quoi ça sert ?
2. Quelles sont les méthodes de sélection de variables les plus importantes et comment fonctionnent-elles ?
3. Comment faire pour mettre en œuvre une méthode de sélection de variable sur de vraies données ?
4. Proposez un démonstrateur, une démonstration sous matlab construite pour convaincre votre patron du bien fondé de vos idées sur la question.

**Exercice 2****L'épreuve du réel****12 points**

Le but de cet exercice est d'écrire la fonction matlab permettant de prédire l'étiquette (0 ou 1) une entrée future  $x_f$ , un vecteur de 10 composants (dont certaines sont inutiles). Pour ce faire, nous disposons des données présentes dans le fichier `lesData` disponible sur moodle.

- les entrées  $X$
- les étiquettes  $y$

```
>> lesData
>> whos
  Name      Size      Bytes  Class  Attributes
  X         93141x10    1490256  single
  y         93141x1     745128  double
>>
```

Donner un code matlab permettant de résoudre ce problème. On insistera sur la méthodologie mise en œuvre